# On Rödl's Theorem for Cographs

Lior Gishboliner[a] Asaf Shapira[b]

Submitted: Jul 7, 2023; Accepted: Sep 21, 2023; Published: Oct 20, 2023



**Abstract**

A theorem of Rödl states that for every fixed $F$ and $\varepsilon > 0$ there is $\delta = \delta_F(\varepsilon)$ so that every induced $F$-free graph contains a vertex set of size $\delta n$ whose edge density is either at most $\varepsilon$ or at least $1 - \varepsilon$. Rödl's proof relied on the regularity lemma, hence it supplied only a tower-type bound for $\delta$. Fox and Sudakov conjectured that $\delta$ can be made polynomial in $\varepsilon$, and a recent result of Fox, Nguyen, Scott and Seymour shows that this conjecture holds when $F = P_4$. In fact, they show that the same conclusion holds even if $G$ contains few copies of $P_4$. In this note we give a short proof of a more general statement.

**Mathematics Subject Classifications:** 05C35, 05C55

## 1 Introduction

Our investigation here is related to two of the most well studied problems in extremal graph theory. A graph-family $\mathcal{F}$ has the *Erdős-Hajnal property* if there is $c = c(\mathcal{F}) > 0$ such that every $n$-vertex induced $\mathcal{F}$-free graph has a clique of independent set of size at least $cn^c$. The famous Erdős-Hajnal conjecture [4] states that every non-empty family of graphs $\mathcal{F}$ has the Erdős-Hajnal property. A variant of the Erdős-Hajnal conjecture was obtained by Rödl [10], who proved that if $G$ is induced $F$-free then for every $\varepsilon > 0$, $G$ contains a set of vertices of size at least $\delta_F(\varepsilon) \cdot n$ whose edge density is either at most $\varepsilon$ or at least $1 - \varepsilon$ (we will henceforth call such sets $\varepsilon$-homogenous). Rödl's proof relied on Szemerédi's regularity lemma, and thus supplied very weak tower-type bounds for $\delta_F(\varepsilon)$. Fox and Sudakov [6] obtained a quantitative improvement over Rödl's proof by showing that one can take $\delta_F(\varepsilon) = \varepsilon^{O_F(\log 1/\varepsilon)}$. This was further improved to $\delta_F(\varepsilon) = \varepsilon^{O_F(\frac{\log 1/\varepsilon}{\log\log 1/\varepsilon})}$ in a recent work of Bucić, Nguyen, Scott and Seymour [3]. Fox and Sudakov [6] made the conjecture that one can take $\delta = \varepsilon^{O_F(1)}$ and noted that a proof of this conjecture would also resolve the Erdős-Hajnal conjecture.

Motivated by Nikiforov's [9] strengthening of Rödl's theorem, Fox, Nguyen, Scott and Seymour [5] introduced the following variant of the conjecture raised in [6]. Let us say

[a]Department of Mathematics, ETH, Zürich, Switzerland (lior.gishboliner@math.ethz.ch).
[b]School of Mathematics, Tel Aviv University, Tel Aviv 69978, Israel (asafico@tau.ac.il).

that a graph $F$ on $f$ vertices is *viral* if for every $\varepsilon > 0$, there is $\delta = \varepsilon^{O_F(1)}$ so that every graph $G$ that contains at most $\delta n^f$ induced copies of $F$ must contain an $\varepsilon$-homogenous set of size $\delta n$. The main result of [5] was that $P_4$, the path on 4 vertices, is viral. Our aim in this paper is to give a very short proof of this result. In fact, we prove the following much more general statement.

**Theorem 1.** *Let $\mathcal{F}$ be a finite graph family which contains a bipartite, a co-bipartite and a split graph. Suppose that $\mathcal{F}$ satisfies the Erdős-Hajnal property. Then there is $C = C(\mathcal{F})$ such that for every $\varepsilon \in (0, 1/2)$ and for every graph $G$ on $n \geqslant \varepsilon^{-C}$ vertices, if $G$ contains fewer than $\varepsilon^C n^{|V(F)|}$ induced copies of $F$ for every $F \in \mathcal{F}$, then $G$ contains an $\varepsilon$-homogenous set $X$ of size $|X| \geqslant \varepsilon^C n$.*

Since $P_4$ is bipartite, co-bipartite and split, and since it is well known that the Erdős-Hajnal conjecture holds for $P_4$, the fact that $P_4$ is viral follows immediately from Theorem 1.

One of the tools used in [5] is the polynomial removal lemma for $P_4$ of [1]. While our proof of Theorem 1 is inspired by an alternative proof of this result in [7], our proof here is much simpler. It is in fact very similar to the Regularity+Turán+Ramsey proof technique that was first introduced in [10] and later used in numerous works applying the regularity method. The key differences which give the improved polynomial bound are that we replace the application of Szemerédi's regularity lemma with an application of the Alon-Fischer-Newman regularity lemma [2], and that we replace the application of Ramsey's theorem with an application of the (assumed) Erdős-Hajnal property.

## 2 Proof of Theorem 1

We will need the following lemma.

**Lemma 2.** *Let $\mathcal{F}$ be a family of graphs which contains a bipartite graph, a co-bipartite graph, and a split graph. Then there is $d = d(\mathcal{F})$ such that for every $\gamma \in (0, 1/2)$, the following holds. For every $k_0 \geqslant 1$ and every graph $G$ on $n \geqslant k_0 \cdot \gamma^{-d}$ vertices, if $G$ contains fewer than $\gamma^d n^{|V(F)|}$ induced copies of $F$ for every $F \in \mathcal{F}$, then $G$ has an equipartition into $k$ parts $V_1, \dots, V_k$, where $k_0 \leqslant k \leqslant k_0 \cdot \gamma^{-d}$, such that for all but $\gamma\binom{k}{2}$ of the pairs $1 \leqslant i < j \leqslant k$ it holds that $d(V_i, V_j) \geqslant 1 - \gamma$ or $d(V_i, V_j) \leqslant \gamma$.*

Lemma 2 was proved in [7], and it follows by combining Lemmas 2.2 and 2.8 in that paper. Indeed, one first uses the fact that $\mathcal{F}$ contains a bipartite, a co-bipartite and a split graph to show that there is a fixed-size bipartite graph $H$ with sides $X, Y$ such that if $G$ is induced $\mathcal{F}$-free then $G$ has no *induced bipartite copy of* $H$, namely, an induced copy of the edges between $X$ and $Y$ (see Lemma 2.2 in [7]). We note that the existence of such an $H$ is equivalent to the statement that induced $\mathcal{F}$-free graphs have bounded VC-dimension. One then uses the conditional regularity lemma of Alon, Fischer and Newman [2] (see also [8]) to show that either $G$ contains many induced bipartite copies of $H$ (in which case $G$ contains many induced copies of some $F \in \mathcal{F}$) or $G$ has a partition as asserted in Lemma 2 (see Lemma 2.8 in [7] for the details).

*Proof of Theorem 1.* Let $f := \max_{F \in \mathcal{F}} |V(F)|$ and $c := c(\mathcal{F})$, namely, $c$ is the constant attesting that $\mathcal{F}$ has the Erdős-Hajnal property. Let $d = d(\mathcal{F})$ be given by Lemma 2. Let $\varepsilon \in (0, 1/2)$. The required constant $C = C(\mathcal{F})$ will be given implicitly by the proof; it will only depend in $f, c, d$. Set

$$\gamma := \min\left\{ \frac{1}{f^2}, \frac{\varepsilon}{4}, \frac{1}{2}\left(\frac{c\varepsilon}{4}\right)^{1/c} \right\}$$

and $k_0 := \lceil 1/\gamma \rceil$. Let $G$ be a graph on $n \geqslant k_0 \cdot \gamma^{-d}$ vertices. We may assume that $G$ contains fewer than $\gamma^d n^{|V(F)|}$ induced copies of $F$ for every $F \in \mathcal{F}$, because $\gamma^d \geqslant \varepsilon^C$ holds provided that $C$ is large enough in terms of $f, c, d$. So we may apply Lemma 2 to obtain a partition $V_1, \dots, V_k$ of $G$ with the properties stated in that lemma. Note that $k \leqslant k_0 \cdot \gamma^{-d} \leqslant \gamma^{-d-1}$. Define an auxiliary graph $R'$ on $[k]$ where $ij \in E(R')$ if $d(V_i, V_j) \geqslant 1 - \gamma$ or $d(V_i, V_j) \leqslant \gamma$. Lemma 2 guarantees that $e(R') \geqslant (1-\gamma)\binom{k}{2} \geqslant (1-2\gamma)\frac{k^2}{2}$, using that $k \geqslant k_0 \geqslant 1/\gamma$. By Turán's theorem, $R'$ contains a clique $A$ of size $|A| = \lceil \frac{1}{2\gamma} \rceil$. Now define an auxiliary graph $R$ on $A$ where $ij$ is an edge if $d(V_i, V_j) \geqslant 1 - \gamma$ and $ij$ is a non-edge if $d(V_i, V_j) \leqslant \gamma$.

**Case 1: $\boldsymbol{R}$ contains an induced copy of some $\boldsymbol{F \in \mathcal{F}}$.** Let $i_1, \dots, i_m \in A$ be the vertices of such a copy; so $m = |V(F)| \leqslant f$. Sample $v_{i_j} \in V_{i_j}$ uniformly at random and independently, $j = 1, \dots, m$. By the definition of $R$, for each $i_j i_\ell \in E(F)$ we have $d(V_{i_j}, V_{i_\ell}) \geqslant 1 - \gamma$, and for each $i_j i_\ell \notin E(F)$ we have $d(V_{i_j}, V_{i_\ell}) \leqslant \gamma$. By the union bound, the probability that $v_{i_1}, \dots, v_{i_m}$ **do not** span an induced copy of $F$ is at most $\binom{m}{2}\varepsilon \leqslant \binom{f}{2}\gamma < \frac{1}{2}$, using that $\gamma < 1/f^2$. It follows that $G$ has at least $\frac{1}{2}|V_{i_1}| \cdot \dots \cdot |V_{i_m}| = \frac{1}{2}(n/k)^m \geqslant \frac{1}{2}\gamma^{(d+1)m} \cdot n^{|V(F)|} \geqslant \varepsilon^C n^{|V(F)|}$ induced copies of $F$, where the last inequality holds provided that $C \gg f, c, d$. This completes the proof in Case 1.

**Case 2: $\boldsymbol{R}$ is induced $\boldsymbol{\mathcal{F}}$-free.** By the choice of $c = c(\mathcal{F})$, the graph $R$ contains a clique or independent set $B \subseteq A$ of size $|B| \geqslant c|A|^c \geqslant c(\frac{1}{2\gamma})^c \geqslant 4/\varepsilon$, using our choice of $\gamma$. Suppose without loss of generality that $B$ is an independent set, and write $B = \{i_1, \dots, i_t\}$. For every $1 \leqslant j < \ell \leqslant t$, we have $d(V_{i_j}, V_{i_\ell}) \leqslant \gamma \leqslant \frac{\varepsilon}{4}$. Also, the number of edges which are contained in one of the sets $V_{i_1}, \dots, V_{i_t}$ is at most $t \cdot \binom{n/k}{2} \leqslant \frac{tn^2}{2k^2} \leqslant \frac{\varepsilon t^2 n^2}{8k^2}$, using that $t \geqslant 4/\varepsilon$. Hence, setting $X = V_{i_1} \cup \dots \cup V_{i_t}$, we have $|X| = t\frac{n}{k} \geqslant \gamma^{d+1} n \geqslant \varepsilon^C n$ and

$$e(X) \leqslant \frac{\varepsilon t^2 n^2}{8k^2} + \binom{t}{2}\frac{\varepsilon}{4} \cdot \frac{n^2}{k^2} \leqslant \frac{\varepsilon t^2 n^2}{4k^2}.$$

As $\binom{|X|}{2} = \binom{tn/k}{2} \geqslant \frac{t^2 n^2}{4k^2}$, we have that

$$d(X) = \frac{e(X)}{\binom{X}{2}} \leqslant \frac{\varepsilon t^2 n^2/(4k^2)}{t^2 n^2/(4k^2)} \leqslant \varepsilon,$$

as required. □

**Acknowledgements**

The first author is supported by SNSF grant 200021_196965. The second author is supported in part by ERC Consolidator Grant 863438 and NSF-BSF Grant 20196.

## References

[1] N. Alon and J. Fox, Easily testable graph properties, Combin. Probab. Comput. 24 (2015), 646–657.

[2] N. Alon, E. Fischer, and I. Newman, Efficient testing of bipartite graphs for forbidden induced subgraphs, SIAM J. Comput. 37 (2007), 959976.

[3] M. Bucić, T. Nguyen, A. Scott and P. Seymour, Induced subgraph density I. A loglog step towards Erdős-Hajnal. arXiv:2301.10147.

[4] P. Erdős and A. Hajnal, Ramsey-type theorems, Discrete Applied Mathematics 25 (1989), 3752.

[5] J. Fox, T. Nguyen, A. Scott and P. Seymour, Induced subgraph density. II. Sparse and dense sets in cographs, arXiv:2307.00801.

[6] J. Fox and B. Sudakov, Induced Ramsey-type theorems, Advances in Mathematics 219 (2008), 17711800.

[7] L. Gishboliner and A. Shapira, Removal lemmas with polynomial bounds, International Math Research Notices (IMRN) (2021), 1440914444.

[8] L. Lovász and B. Szegedy, Regularity partitions and the topology of graphons. In: An Irregular Mind, Imre Bárány, Jozsef Solymosi, and Gábor Sági editors, Bolyai Society Mathematical Studies 21 (2010), 415446.

[9] V. Nikiforov, Edge distribution of graphs with few copies of a given graph, Combin. Probab. Comput. 15 (2006), 895902.

[10] V. Rödl, On universality of graphs with uniformly distributed edges, Discrete Mathematics 59 (1986), 125134.